

Stat 334 Final Project: Appendix
Cameron Stivers and Beau McAndrew
California Polytechnic State University - Stat 334, Winter 2022

Model 1

lm(LenStay ~ Age + InfRisk + Culture + XRay + School + Region + Beds + Census + Nurses + Service, data = hospital)

Table 1A

Call:
lm(formula = LenStay ~ Age + InfRisk + Culture + XRay + School + Region + Beds + Census + Nurses + Services, data = hospital)

Residuals:
Min 1Q Median 3Q Max
-2.3048 -0.6608 -0.0272 0.5862 6.3001

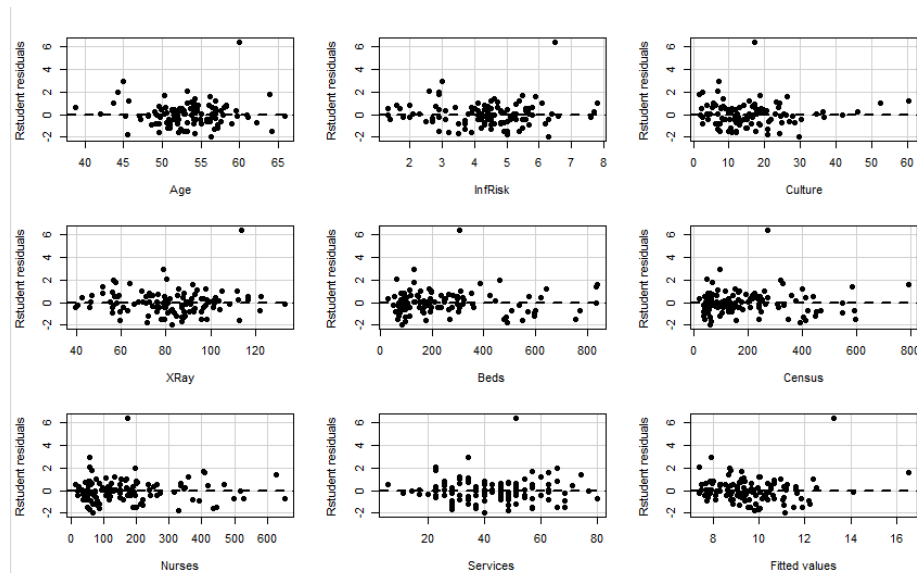
Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.242682	1.612956	1.390	0.167489
Age	0.079922	0.028266	2.827	0.005668 **
InfRisk	0.439665	0.127298	3.454	0.000812 ***
Culture	0.005546	0.015982	0.347	0.729299
XRay	0.012688	0.007147	1.775	0.078892 .
SchoolYes	0.266644	0.441089	0.605	0.546872
RegionNE	0.812966	0.351406	2.313	0.022744 *
Regions	-0.345311	0.313506	-1.101	0.273345
RegionW	-1.067594	0.398736	-2.677	0.008672 **
Beds	-0.004851	0.003603	-1.346	0.181224
Census	0.015182	0.004424	3.432	0.000872 ***
Nurses	-0.005891	0.002218	-2.656	0.009203 **
Services	-0.012179	0.013774	-0.884	0.378698

signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.231 on 100 degrees of freedom
Multiple R-squared: 0.6299, Adjusted R-squared: 0.5855
F-statistic: 14.18 on 12 and 100 DF, p-value: < 2.2e-16

Plot 1B



Plot 1C

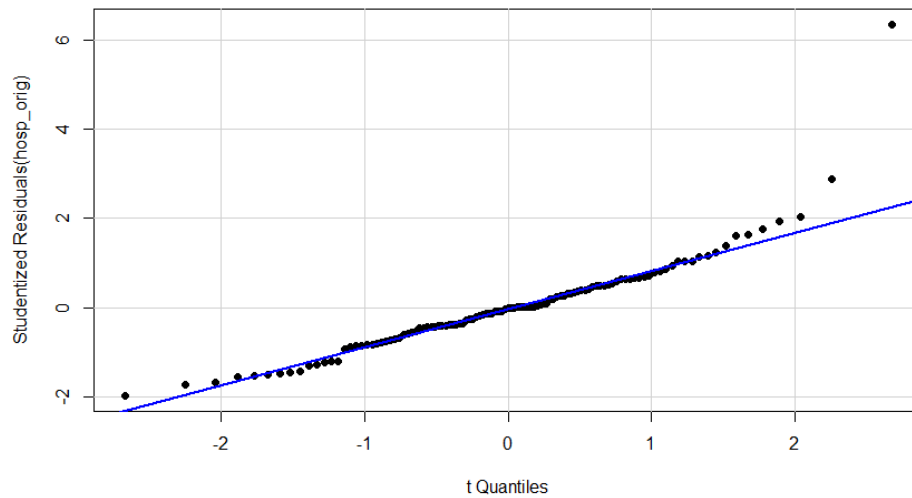
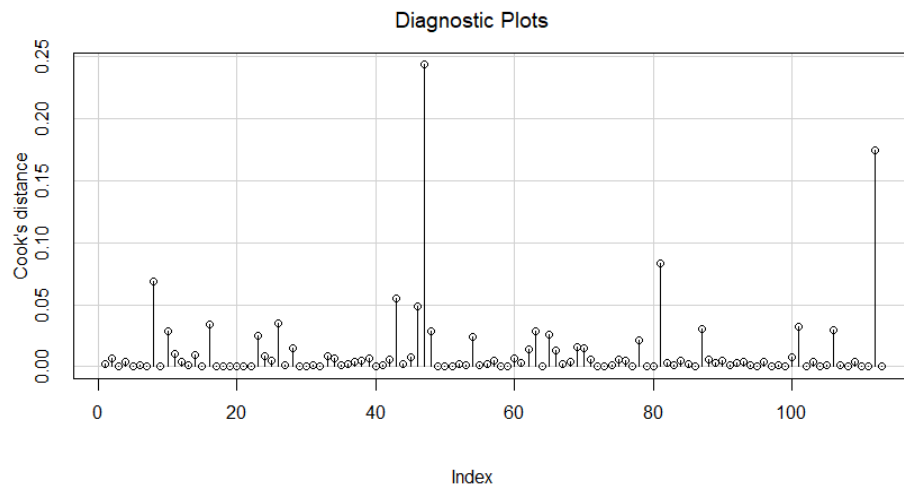


Table 1D

	Variables	Tolerance	VIF
1	Age	0.85021570	1.176172
2	InfRisk	0.46410296	2.154694
3	Culture	0.50542828	1.978520
4	XRay	0.70608230	1.416265
5	SchoolYes	0.53898638	1.855334
6	RegionNE	0.58231265	1.717291
7	Regions	0.61921830	1.614939
8	RegionW	0.69356705	1.441822
9	Beds	0.02801183	35.699204
10	Census	0.02923000	34.211423
11	Nurses	0.14173293	7.055523
12	Services	0.30846950	3.241812

Plot 1E



Model 2

lm(log(LenStay) ~ Age + InfRisk + log(Culture) + XRay + School + Region + bedprop + Census + nurseprop + Services, data = hospital)

Table 2A

```
call:
lm(formula = log(LenStay) ~ Age + InfRisk + log(Culture) + XRay +
    School + Region + bedprop + Census + nurseprop + Services,
    data = hospital)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.23201	-0.06240	-0.00809	0.06161	0.39575

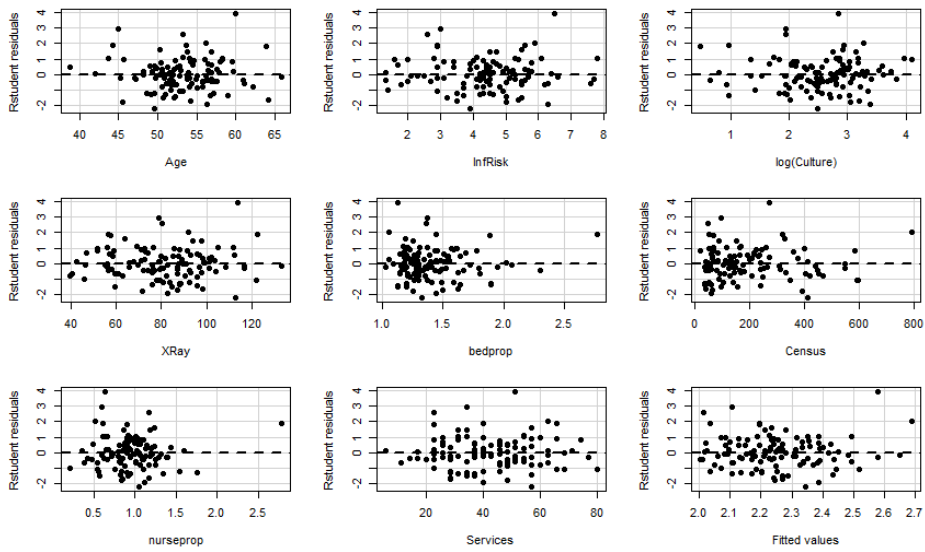
Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	1.7226454	0.1658024	10.390	< 2e-16	***
Age	0.0066194	0.0026597	2.489	0.01447	*
InfRisk	0.0533452	0.0119114	4.479	2e-05	***
log(Culture)	-0.0230057	0.0235087	-0.979	0.33014	
XRay	0.0014337	0.0006759	2.121	0.03638	*
SchoolYes	0.0454738	0.0390386	1.165	0.24685	
RegionNE	0.0797853	0.0318653	2.504	0.01390	*
Regions	-0.0373107	0.0292232	-1.277	0.20465	
RegionW	-0.1255890	0.0386397	-3.250	0.00157	**
bedprop	-0.0362832	0.0525680	-0.690	0.49166	
Census	0.0003303	0.0001334	2.476	0.01497	*
nurseprop	-0.0664807	0.0443164	-1.500	0.13673	
Services	-0.0014225	0.0012253	-1.161	0.24844	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1135 on 100 degrees of freedom
 Multiple R-squared: 0.637, Adjusted R-squared: 0.5935
 F-statistic: 14.62 on 12 and 100 DF, p-value: < 2.2e-16

Plot 2B



Plot 2C

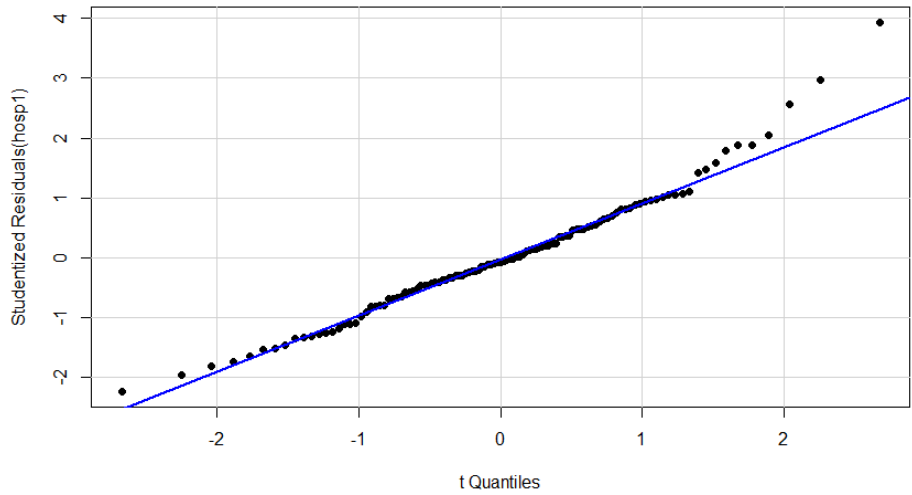


Table 2D

	variables	Tolerance	VIF
1	Age	0.8164531	1.224810
2	InfRisk	0.4506695	2.218921
3	log(Culture)	0.4303468	2.323707
4	XRay	0.6711802	1.489913
5	SchoolYes	0.5850186	1.709347
6	RegionNE	0.6020951	1.660867
7	Regions	0.6059088	1.650413
8	RegionW	0.6279402	1.592508
9	bedprop	0.5788174	1.727660
10	Census	0.2732073	3.660225
11	nurseprop	0.5661638	1.766273
12	Services	0.3314038	3.017467

Model 3

lm(log(LenStay) ~ Age + InfRisk + log(Culture) + XRay + School + Region + bedprop + Census + nurseprop + Services + InfRisk * Census, data = hospital)

Table 3A

```
call:
lm(formula = log(LenStay) ~ Age + InfRisk + log(Culture) + XRay
    School + Region + bedprop + Census + nurseprop + Services +
    InfRisk * Census, data = hospital)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-0.20513 -0.06155 -0.01110  0.06719  0.37681
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.922e+00  1.871e-01  10.271 < 2e-16 ***
Age          4.623e-03  2.771e-03   1.669  0.09837 .
InfRisk      2.773e-02  1.665e-02   1.666  0.09893 .
log(Culture) -2.504e-02  2.311e-02  -1.084  0.28112
XRay         1.543e-03  6.657e-04   2.318  0.02250 *
SchoolYes    2.291e-02  3.974e-02   0.577  0.56552
RegionNE     7.495e-02  3.138e-02   2.389  0.01880 *
Regions     -3.712e-02  2.870e-02  -1.293  0.19895
RegionW     -1.157e-01  3.822e-02  -3.028  0.00314 **
bedprop     -5.259e-02  5.218e-02  -1.008  0.31597
Census      -5.325e-04  4.200e-04  -1.268  0.20785
nurseprop   -5.745e-02  4.372e-02  -1.314  0.19191
Services    -4.412e-04  1.286e-03  -0.343  0.73228
InfRisk:Census 1.699e-04  7.857e-05   2.162  0.03302 *
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.1114 on 99 degrees of freedom
Multiple R-squared:  0.6534,    Adjusted R-squared:  0.6079
F-statistic: 14.36 on 13 and 99 DF,  p-value: < 2.2e-16
```

Model 4

**lm(log(LenStay) ~ Age + Region + XRay + nurseprop + Census + InfRisk +
InfRisk * Census, data = hospital)**

Table 4A

```
> best.hosp <- ols_step_forward_aic(hosp_int)
> best.hosp[["predictors"]]
[1] "InfRisk:Census" "Region"          "Census"          "XRay"            "nurseprop"       "Age"
[7] "InfRisk"
```

Table 4B

```
Call:
lm(formula = log(LenStay) ~ Age + Region + XRay + nurseprop +
    Census + InfRisk + InfRisk * Census, data = hospital)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.21879	-0.06706	-0.01041	0.06883	0.38169

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.815e+00	1.611e-01	11.264	< 2e-16 ***
Age	4.933e-03	2.520e-03	1.957	0.05303 .
RegionNE	6.887e-02	2.971e-02	2.319	0.02240 *
RegionS	-3.895e-02	2.770e-02	-1.406	0.16262
RegionW	-1.180e-01	3.671e-02	-3.215	0.00174 **
XRay	1.391e-03	6.379e-04	2.180	0.03150 *
nurseprop	-8.642e-02	3.785e-02	-2.283	0.02447 *
Census	-5.417e-04	3.462e-04	-1.565	0.12074
InfRisk	2.236e-02	1.402e-02	1.595	0.11385
Census:InfRisk	1.755e-04	7.022e-05	2.500	0.01401 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1106 on 103 degrees of freedom
Multiple R-squared: 0.6446, Adjusted R-squared: 0.6136
F-statistic: 20.76 on 9 and 103 DF, p-value: < 2.2e-16

Model 5

**lm(log(LenStay) ~ Region + XRay + nurseprop + Census + InfRisk +
InfRisk * Census, data = hospital)**

Table 5A

```
Call:
lm(formula = log(LenStay) ~ Region + XRay + nurseprop + Census +
    InfRisk + InfRisk * Census, data = hospital)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.25683	-0.06478	-0.00880	0.06682	0.40453

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	2.098e+00	7.177e-02	29.233	< 2e-16	***
RegionNE	7.929e-02	2.962e-02	2.677	0.00864	**
Regions	-3.151e-02	2.780e-02	-1.133	0.25966	
RegionW	-1.107e-01	3.702e-02	-2.991	0.00348	**
XRay	1.356e-03	6.462e-04	2.098	0.03833	*
nurseprop	-8.487e-02	3.835e-02	-2.213	0.02909	*
Census	-7.495e-04	3.340e-04	-2.244	0.02695	*
InfRisk	1.699e-02	1.393e-02	1.219	0.22549	
Census:InfRisk	2.170e-04	6.785e-05	3.198	0.00183	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1121 on 104 degrees of freedom
Multiple R-squared: 0.6314, Adjusted R-squared: 0.603
F-statistic: 22.27 on 8 and 104 DF, p-value: < 2.2e-16

Table 5B

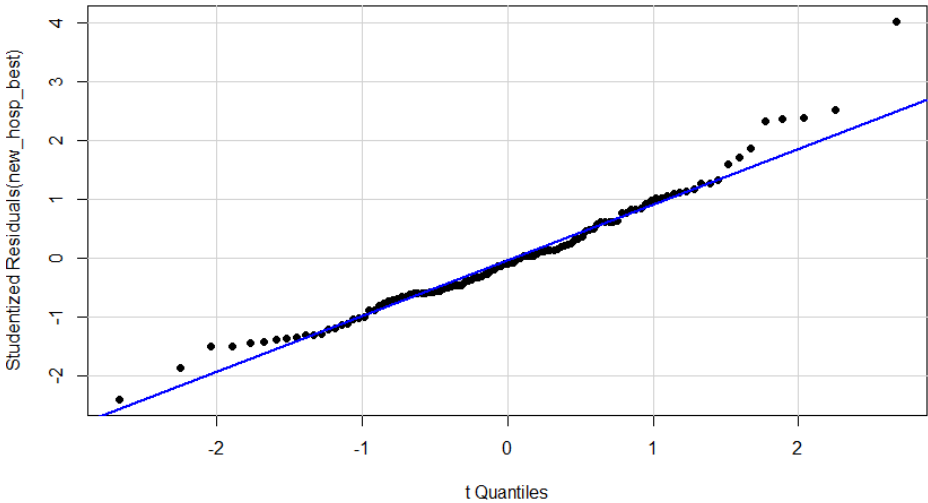
Anova Table (Type II tests)

Response: log2LenStay

	Sum Sq	Df	F value	Pr(>F)	
Region	0.71843	3	9.1511	1.986e-05	***
XRay	0.11519	1	4.4017	0.0383259	*
nurseprop	0.12815	1	4.8969	0.0290905	*
Census	0.35495	1	13.5637	0.0003679	***
InfRisk	0.61056	1	23.3310	4.714e-06	***
Census:InfRisk	0.26764	1	10.2272	0.0018340	**
Residuals	2.72162	104			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Plot 5C



Plot 5D

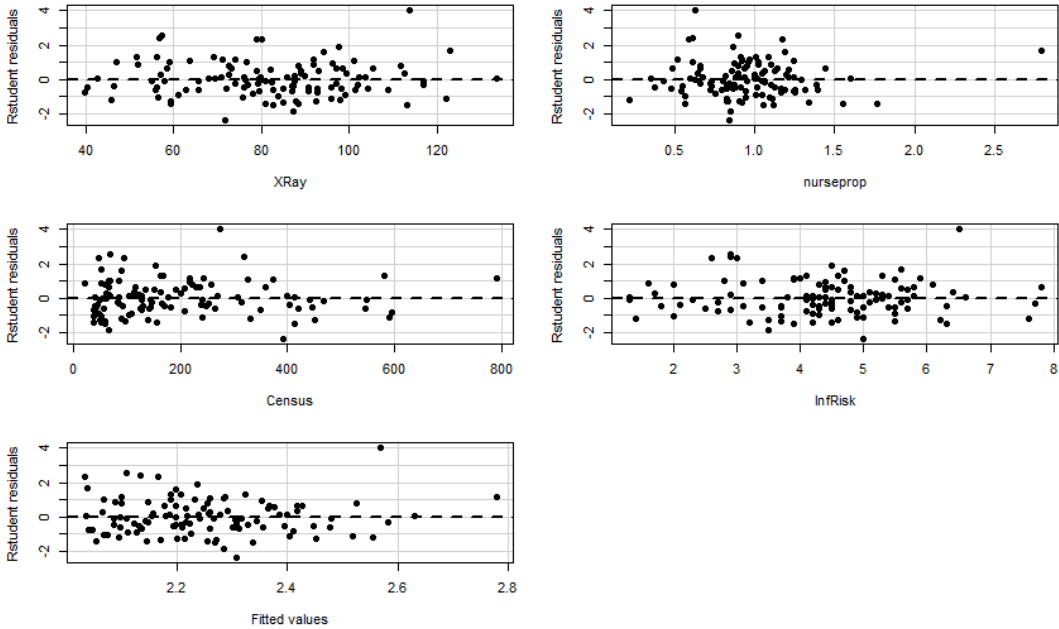


Table 5E

	Variables	Tolerance	VIF
1	RegionNE	0.68038495	1.469756
2	Regions	0.65354166	1.530124
3	RegionW	0.66814518	1.496681
4	XRay	0.71694015	1.394817
5	nurseprop	0.73809612	1.354837
6	Census	0.04256648	23.492664
7	InfRisk	0.32153156	3.110115
8	Census:InfRisk	0.03398529	29.424497

Plot 5F

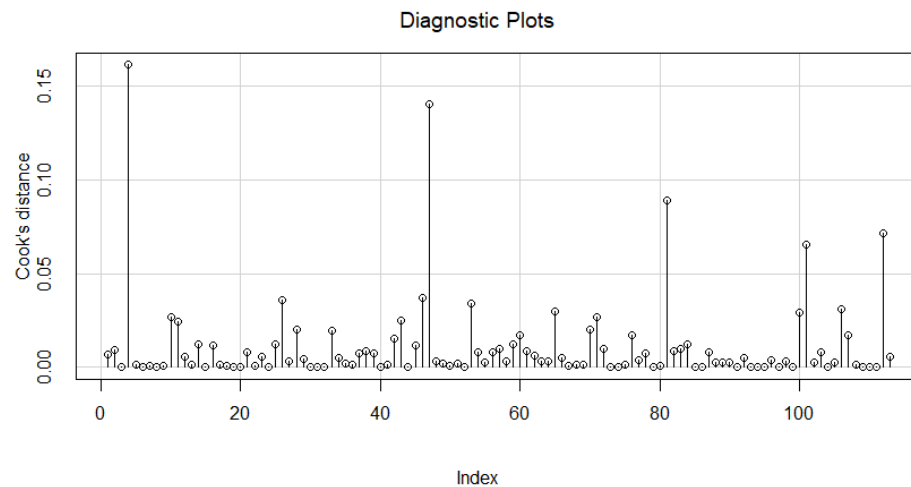


Table 5G

Analysis of Variance Table

Model 1: $\log(\text{LenStay}) \sim \text{Region} + \text{XRay} + \text{nurseprop} + \text{Census} + \text{InfRisk} + \text{InfRisk} * \text{Census}$

Model 2: $\log(\text{LenStay}) \sim \text{Age} + \text{InfRisk} + \log(\text{Culture}) + \text{XRay} + \text{School} + \text{Region} + \text{bedprop} + \text{Census} + \text{nurseprop} + \text{Services} + \text{InfRisk} * \text{Census}$

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	104	1.3076				
2	99	1.2296	5	0.078016	1.2563	0.289